Conceptual short-term memory: a missing part of the mind?

Forthcoming in the *Journal of Consciousness Studies*

Abstract

In debates in philosophy and cognitive science concerning short-term memory mechanisms and perceptual experience, most discussion has focused on the working memory and the various forms of sensory memory such as iconic memory. In this paper, I present a summary of some evidence for a proposed further form of memory termed Conceptual Short-Term Memory. I go on to outline some of the ways in which this additional distinctive sort of short-term memory might be of relevance to ongoing philosophical debates, specifically in relation to questions about high-level perceptual phenomenology, the relationship between consciousness and reportability, and the boundary between cognition and perception. I conclude that Conceptual Short-Term Memory offers a promising new direction of research and philosophical investigation.

Conceptual short-term memory: a missing part of the mind?

When we engage in almost any perceptual activity – recognizing a face, listening out for a phone-call, or simply taking in a sunset – perceptual information must be briefly stored and processed. Most such information is almost immediately lost, with only a small fraction of it leaving traces that can be later recovered. Despite its ephemerality, this kind of fleeting perceptual representation constitutes a major part of our mental lives. Thus in attempting to develop an empirically grounded account of perception and conscious experience, it is crucial to engage with scientific theories of the kinds of short-term memory mechanisms that underlie our moment-to-moment retention of information about the world.

While memory research has made significant progress over the last sixty years, there are still a number of competing models and many unresolved questions. Nevertheless, within this complex literature, two distinct forms of memory are fairly well-defined, namely high-capacity short-term sensory stores (Sperling, 1960; Landman et al., 2003), and a more capacity-limited cognitive mechanism sometimes referred to as working memory (Cowan 2001, Baddeley 2000). However, psychologists have also recognized that these mechanisms are probably not exhaustive of short-term memory processes (Cowan, 1988). In this paper, I will present a selection of recent empirical evidence in support of one such further posited short-term memory mechanism termed Conceptual Short Term Memory (CSTM) which I believe may be particularly promising for certain philosophical debates.

I begin in the first part of this paper by providing a brief overview of some current psychological literature on sensory memory and working memory. I go on in the second part of the article to describe the main empirical data supporting the existence of CSTM. The third part of the paper surveys some possible applications of CSTM to ongoing debates in philosophy of mind and cognitive science, specifically relating to high-level perceptual content, the relationship between consciousness and reportability, and the perception-cognition boundary. I conclude by suggesting that, if empirical data continues to support the existence of a mechanism broadly equivalent to CSTM, it may constitute a fruitful addition to theoretical and philosophical debates about the mind.

1. Sensory memory and working memory

Before spelling out the CSTM hypothesis in detail, it is worth briefly expanding upon two of the best defined and most widely accepted forms of short-term processing in current models, namely *sensory stores* and limited capacity *working memory*. The basic distinction between these forms of short-term memory draws upon the work of Sperling (1960) and Miller (1956), and was spelled out explicitly in the important Atkinson-Shiffrin memory model (Atkinson & Shiffrin, 1968) that divides memory systems into sensory registers, a short term cognitive store, and long-term memory.

The Atkinson-Shiffrin model has since been superseded by more complex accounts, notably the frameworks of Alan Baddeley (Baddeley & Hitch, 1974; Baddeley, 2001) and Nelson Cowan (1988, 2001). However, while Baddeley and Cowan's models of short-term memory differ from one another in important respects, they, together with most other theorists in memory research, nonetheless still allow for a distinction between brief high-capacity sensory stores and a more robust but capacity-limited cognitive store that Baddeley terms 'the episodic buffer' and Cowan calls 'the focus of attention'. Following predominant patterns of usage in contemporary scientific and philosophical literature, I will refer to this latter multi-modal capacity-limited informational store as working memory'

Though it has previously been suggested that these two stores are probably not the only forms of short-term memory (Cowan, 1988), the distinction between them has nonetheless been highly influential in models of perception and cognition, especially in the philosophical literature (e.g., Dretske, 1981: Ch.11; Block, 2007), and thus is worth spelling out in more detail.

Consider sensory memory first. This category encompasses several distinct forms of memory that can be characterized by their relatively high capacities, brief durations, and susceptibility to disruption by the presentation of new information in the relevant sensory modality. Important early evidence for the existence of such sensory stores came from Sperling's work on partial report paradigms, which demonstrated the existence of a form of sensory memory dubbed iconic memory. In his original experiment, subjects were briefly shown a 3x4 matrix of alphanumeric characters (Sperling, 1960). After the stimulus was removed, if subjects were immediately cued to report on any given row, they were able to successfully report almost all the contents of that one row (3-4 items), though not other rows. This suggests that subjects have brief access (for up to 300ms) to a fragile memory store that encoded 9-12 items from the initial stimulus.

This is now good evidence for sensory memory in other modalities, including audition (Darwin, Turvey, & Crowder, 1972) and tactile sensation (Harris et al., 2002). Echoic memory, for example, serves a similar function to iconic memory in the auditory domain, and like iconic memory has a very large capacity, but persists for longer durations of up to 2 seconds. There is also some evidence for a second stage of visual sensory memory, dubbed Fragile Visual Short-Term Memory (fVSTM). This has been claimed to persist significantly longer than iconic memory, and to have an even higher capacity than Sperling's iconic memory, with an upper limit of at least 32 objects (Sligte, Scholte, & Lamme, 2008).

There are two further features of sensory memory that are worth briefly dwelling on. The first is that, at least in the case of visual sensory memory, sensory memory is readily disrupted by further stimulation in the relevant sensory modality, such as a pattern mask in the case of iconic memory and fVSTM (Sligte, Scholte, & Lamme, 2008; Averbach & Coriell, 1961; Saults & Cowan, 2007, show overwriting effects in echoic memory). The second is that while visual sensory memory allows in some cases for feature binding (Landman, Spekreijse, & Lamme, 2003), current evidence seems to indicate that it encodes objects just in respect of fairly low-level sensory properties such as color and shape, rather than higher-level semantic properties such as conceptual identity. This is suggested, for example, by the discovery that subjects do not exhibit partial report superiority when cued to report just on items belonging to a given semantic category, such as 'numbers' or 'letters' (Sperling, 1960; von Wright, 1970).

All of these features help distinguish sensory memory from working memory. Whereas sensory memory is rich but fragile, working memory is a robust, flexible, but strictly capacity-limited store recruited for complex attentionally-demanding tasks, such as memorizing lists of unfamiliar numbers or words or picking out a previously seen color patch from an array (Cowan, 2001; Luck & Vogel, 1997). It is capable of encoding information not just in terms of its low-level sensory properties, but also in terms of high-level concepts.

Early evidence for some such capacity-limited form of central short-term memory was presented by Miller (1956), who claimed that many forms of short-term recall were constrained by a processing limit of around seven items at once. Subsequent research suggested that matters were considerably more complicated, however. First, people are frequently able to remember much more than seven items at a time via the use of chunking; that is, combining different words, numbers, or even visual spatial items into a single representation for the purposes of memory storage. A second complicating factor in assessing working memory capacity comes from the fact that sensory processes can be actively recruited by working memory for retention of information over extended intervals. For example, our ability to remember lists of words over short intervals seems to sometimes involve not just working memory in the strictest sense but also auditory sensory representations that are sustained by active rehearsal (Baddeley, Thomson, & Buchanan, 1975; Baddeley, Lewis & Vallar, 1984).

Nonetheless, when experiments are designed to control for both chunking and rehearsal in sensory memory we find a surprisingly high degree of regularity in people's capacity to store distinct pieces of information, namely a limit of around four items at a time (Cowan, 2001). This holds true regardless of whether the relevant information to be retained is semantic (as in the identity of a digit or the meaning of a word) or tied to a particular sensory modality (as in a color patch or tone).[1]

However, it is worth noting that while the four item limit shows up in a wide range of empirical paradigms, there is some evidence that it may vary depending on the precise content that is being stored (Brady, Störmer, & Alvarez, 2016). Its capacity may also not be best understood in terms of a literal number of finite 'slots' (as sometimes suggested, e.g., Cowan 2001 and Luck & Vogel 1997), but rather in terms of the difficulties faced by the brain in

[1] Note that this is not simply the claim that *attention* is limited to four items. The relationship between attention and working memory remains controversial. See Fougnie (2008) for a review.

recovering signals from ‘noisy’ representations, where noise increases as more and more information is maintained at once (Ma et al. 2014).[2]

2. Conceptual short-term memory

As noted, sensory stores and working memory are probably not the only form of short-term mechanism.[3] Cowan (1988), for example, suggests that there may be a further form of short-term memory subserved by briefly activated portions of long-term memory outside of focal attention, while Ericsson and Kintsch (1995) posit a special form of Long-Term Working Memory to explain durable retention of information during tasks involving disruptions to short-term working memory.

In the remainder of this paper, I wish to focus on one such hypothesized store, Conceptual Short-Term Memory, developed in the work of Mary Potter. Though CSTM shares some common features with other proposed short-term stores, I believe it to be of particular interest and relevance to contemporary theorists given that it draws upon a large body of research, much of it very recent, and its putative features have been spelled out by Potter in some detail (Potter, 2009, 2012). It may also, I suggest, offer important philosophical insights in understanding perceptual experience.

The basic notion of CSTM is that it is a further short-term store that serves as “a mental buffer in which current stimuli and their associated concepts from long term memory… are represented briefly, allowing meaningful patterns or structures to be identified” (Potter, 2009). Unlike sensory memory, it encodes high-level features of objects and scenes and does not seem vulnerable to disruption by new visual inputs. It also has a shorter duration than some of the more robust forms of sensory memory such as fragile visual short-term memory. However, unlike working memory, it has a high capacity and fairly brief duration.[4]

There are a significant number of experiments that have been taken support the existence of some such intermediate store, and a detailed review would lie beyond the scope of this paper.

[2] Another ongoing debate concerns whether working memory is sometimes unconscious. I will assume for the purposes of this paper that working memory is always conscious. See Stein, Kaiser, and Hesselmann (2016) for a review of this ongoing controversy.
[3] Thanks to an anonymous reviewer for stressing this point.
[4] Note that Potter sometimes (e.g., 2012) talks of CSTM as a *type* of working memory, albeit one with quite distinctive features. For the sake of clarity, I will continue to use the term ‘working memory’ just to refer to the more strictly capacity-limited store described earlier.

Accordingly, I will discuss just a few experiments, focusing on those conducted by Potter herself and some others which explicitly appeal to CSTM in explaining their results.

One early piece of suggestive evidence comes from experiments in which subjects seem to have brief but fleeting awareness of large numbers of rapidly sequentially presented stimuli. For example, subjects very briefly presented with the words from a twelve-word *sentence* one at a time can accurately remember every word they see, but cannot do so for a similarly presented list of twelve unrelated words (Potter, Kroll, & Harris, 1980). This suggests that subjects must be able to access the meanings of those words very quickly in order to establish whether there are appropriate semantic and syntactical connections between them. If the relevant connections are present, the sentence can be retained in working memory via chunking mechanisms, but in cases where those connections do not exist, the information is almost immediately lost.

Another experiment of Potter's provided further evidence that this conceptual information is rapidly lost unless consolidated (Potter, Stiefbold, & Moryadas, 1998; see Fig. 1). Subjects saw a sentence rapidly presented one word at a time, each word being displayed for 133ms. At one point in the sentence, subjects briefly saw a pair of words shown simultaneously, only one of which was contextually appropriate. Their task was to pick out the contextually appropriate word and repeat the whole sentence, a task they performed well at. However, they were frequently unable to recall the word whose meaning they had rejected, even though its semantic content must have been in some way accessed in order for subjects to prefer the other word as more contextually appropriate.

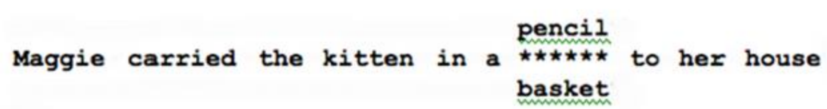

Fig. 1 (from Potter et al. 1998). Subjects saw a series of words like those below, and had to choose the semantically more appropriate of two words before repeating back the whole sentence.

These experiments seem to demonstrate that high-level semantic information about visual stimuli is accessed extremely rapidly but quickly lost. However, it remains possible that this is explicable in terms of well-established forms of memory, such as rapid serial encoding in working memory. More recent work of Potter's is harder to explain away in these terms, however (Potter et al., 2014). In this crucial experiment, subjects were shown 6-12 sequential images (which they had not previously seen) for durations of 13, 27, 53, or 80ms (see Fig. 2). They were given a target description (for example, 'wedding' or 'flowers') 900ms before or 200ms after presentation of the images, and asked to say whether any of the images they saw matched the description. The longer the initial duration of the stimuli, the more likely subjects were to correctly detect a target, but were above chance in all conditions. Despite their above-chance performance on these objective measures, the subjects only reported a rapidly changing, very short sequence of colors and shapes (with the exception of the last picture, which they were not tested on).

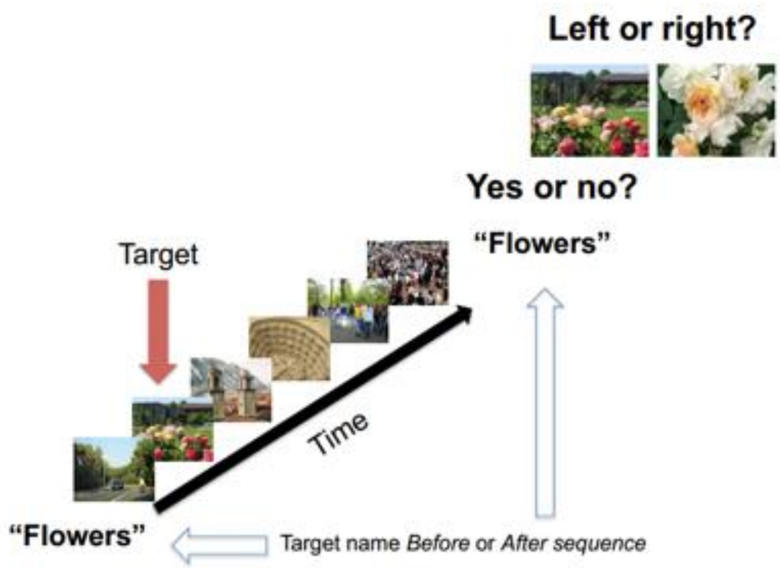

Fig. 2 (from Potter et al. 2014)

Sequences of 6-12 images were rapidly presented. Subjects were given a target description 900ms before or 200ms after presentation and instructed to say whether they saw a picture matching that description. In trials where a picture was presented, they performed a subsequent recognition task.

Some of the trials also incorporated a second task. Having been asked to identify whether they had seen an image matching a given description, subjects were subsequently shown two images matching that description and given a forced choice task in which they had to indicate which of the two images was the one actually presented. They were given this task even on trials where they had not successfully detected the image. Subjects' performance in this second task was closely linked to whether they had made a successful detection of the target under a given description. For subjects who failed to detect the initial stimulus, "their forced choice was near chance, suggesting that the visual features of unidentified pictures were not retained" (Potter et al. 2014: 276).

We can begin to draw a rough picture about what seems to be happening in this experiment. First, in order to be able to reliably perform detection and recognition tasks after stimulus presentation, subjects presumably had to retain information about all or most of the 6-12 presented images. Moreover, in order to be able to use this information to say whether any of the images matched a given target description, subjects must either have already stored semantic information about each image or otherwise be able to rapidly derive it.

So much should be fairly uncontroversial. What is much more open for debate is *which* memory mechanisms underpin subjects' performance. Potter's own assessment is that this result is best explained in terms of a dedicated memory mechanism in the form of CSTM. Building on this suggestion, one natural interpretation is that subjects briefly stored information about all stimuli, encoding both high-level semantic information (so as to be able to detect whether any stimulus fitted a given description) and some lower level information (so as to enable accurate recognition). This was momentarily available, such that, when cued with a target description, subjects were able to give accurate reports and could recognize the relevant image if it was presented to them. In cases where there was no encoding in CSTM, subjects were both unable to detect the image according to its description, and were at chance in recognizing it.

This seems to be the most straightforward explanation of the data, but there are debunking approaches also are worth considering. Specifically, it is not immediately obvious that the results cannot be explained just in terms of sensory memory or ordinary working memory.

Consider first the hypothesis that subjects' performance might be explained in terms of capacity-limited working memory. This might be able to explain subjects' performance in the trials where they are cued in advance: having been cued to look out for a picture matching the description 'wedding', for example, they could rapidly encode semantic information about each image as it is presented, specifically searching for a wedding and discarding information about all other images.

However, this account does not explain subjects' only slightly lower accuracy when they were cued *after* they had seen the images. As noted earlier, working memory has strict capacity limits, well below the twelve images presented in some trials. At most, then, subjects could retain four images in working memory, fewer than the half the images in the array. That might still enable them to perform marginally above chance, but if that was indeed what was responsible for subjects' performance, then we would expect a significant difference in subjects' performance between the 6-item and 12-item trials, since they would go from being able to encode the majority of the stimuli to barely a third of them. In fact, subjects' performance was very similar in the two trials, suggesting the effect is not due to working memory.[5]

An alternative debunking hypothesis might claim that subjects were retaining a *sensory* representation of the stimuli, much as they do in the Sperling Test and other partial report paradigms. More specifically, one might imagine that subjects retain multiple distinct icons corresponding to each of the 6-12 images as they are presented, which they can then conceptualize and 'inspect' after presentation when they are given a target.

This sort of explanation seems unlikely, however. As noted earlier, visual forms of sensory memory like iconic memory and fragile visual short-term memory are both disrupted by the presentation of sequential images in the same location. This is not an issue for the Sperling experiment, since this involves just a single initial stimulus prior to cueing. In Potter's experiment, however, subjects saw *multiple* images one after another in exactly the same location prior to cueing. Any representations in iconic or fragile visual short-term memory would therefore be rapidly 'overwritten' as the sequence was presented.[6]

Additionally, a sensory interpretation would struggle to explain experiments performed by Potter that examined subjects' vulnerability to conceptual 'decoys'. In one such experiment, subjects were shown a sequence of five pictures at 173ms exposures, and then immediately given a test picture and asked whether it was one of the five pictures just presented (Potter, Staub, & O'Connor, 2004). Subjects performed fairly well at this task. However, they were significantly more prone to error when tested on distractors which were similar in semantic content to pictures that had just been presented. These 'decoy' pictures were carefully chosen so as to share the same conceptual gist as one of the pictures shown by Potter without being too visually similar (see Fig. 3).

---

[5] Potter et al. comment that "we can reject the hypothesis that participants could encode only two or three pictures in working memory; otherwise, performance would have fallen more dramatically in Experiment 2, especially in the after condition, in which participants had to retain information about the pictures for later retrieval" (2014: 275).

[6] Note, however, one complication arising in light of follow up work by Maguire & Howe (2016). This showed that more aggressive masking techniques than those used by Potter strongly negatively affected subjects' performance in this paradigm, but only for exposure times of less than 50ms.

This suggests that subjects' ability to remember the pictures was not based purely on the retention of low-level sensory information but involved encoding of the image in terms of its semantic properties. Otherwise, one would not expect the conceptual (but non-pictorial) similarity of the images to have any significant effect on subjects' performance. This is further strong evidence that subjects' performance in Potter's work on CSTM is not simply a matter of sensory short-term memory.

Fig. 3. (from Potter et al. 2005). An example of a target picture ('camel') and its conceptual decoy.

I now wish to present a quite different form of empirical evidence for the CSTM hypothesis which comes from recent work on visual search tasks. Unlike Potter's experiments, most of which involve rapid serial visual presentation, the experiments I will now describe involve presentation of a single array of stimuli, much like the Sperling experiment.[7]

In the first such experiment subjects saw a single grid of objects for brief durations (Moores, Laiti, & Chelazzi, 2003). Subjects looked at a central fixation point, and their task was to assess whether a target stimulus (for example, a motorbike) was present in the array. Moores et al. showed that the presence of semantically-related distractors in the array (e.g., a motorbike helmet) had a negative impact on subjects' reaction times and accuracy in trials where the target was absent. This suggests that subjects were diverted by the presence of pictures that were conceptually similar to their target.

This is not itself surprising, and is compatible with an account cashed out purely in

[7] One might wonder why memory mechanisms are relevant to the following experiments, since the arrays in question remain in view of subjects at all times. However, any short-term processing of information, including perceptual encoding of semantic properties, presumes some kind of underlying short-term storage mechanisms, and I can see no reason see why these mechanisms would be different in cases where the stimulus was still visible. Indeed, in one of the studies (Belke et al., 2008) the experimenters specifically take their data to support the CSTM hypothesis.

terms of working memory; it is possible, for example, that as subjects scanned the array they were momentarily sidetracked by the semantically related items, and had to pause to assess whether they fit the target description. What is much more surprising is that in a majority of trials where the target was absent but a semantically related distractor was present, subjects fixated the distractor *before all other stimuli*. In other words, subjects' looking behavior already seemed sensitive to the semantic properties of the items in the array even before they had the chance to visually fixate them.

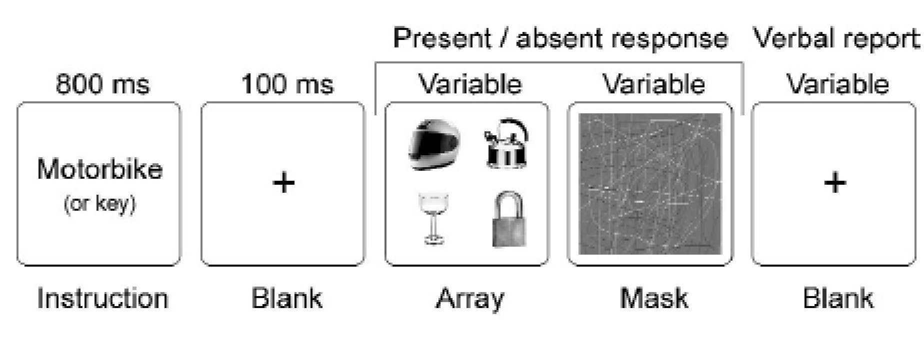

Figure 4 (from Moores et al. 2003)

Subjects are given a verbal target, then fixate a central point. They are presented with an array of objects for brief durations (average 126ms) and asked whether the target is present. Eye-tracking showed that in a majority of trials subjects initially fixated the target object if present or semantically related distractors if the target was absent.

This is difficult to explain just in terms of iconic memory, since, as noted, sensory forms of memory do not seem to be sensitive to the semantic properties of stimuli (subjects in the Sperling task could not be trained to report just on letters or just on numbers, for example). However as noted above, subjects' eye movement suggest that items in the array have been in some way encoded in respect of their semantic properties.

Nonetheless, since there are only four items present, this result could be explained in terms of working memory; for example, subjects might be engaging in a kind of rapid covert attention to all the items in the array, quickly extracting semantic information about all four of them. This interpretation looks much less plausible, however, in light of a later similar experiment (Belke et al., 2008). This experiment used a broadly similar methodology, but varied the number of objects in the trial, with arrays containing up to eight items. Crucially, this increase in the size of the array did not affect the likelihood that subjects initially fixated the target or semantically-related distractor object; instead, subjects immediately looked at the target or distractor object in a majority of trials.

This result clearly runs counter to the hypothesis that the experiments can be explained just in terms of working memory. If subjects were indeed just using working memory together with some form of covert attention to access the semantic properties of each item to begin with, then the size of the array should make a significant difference to how frequently and reliably subjects directed their initial eye movements to semantically relevant targets or distractors. Instead, subjects initially looked at semantically relevant objects with equal frequency in trials with 4 objects and trials with 8 objects. This suggests that the rapid semantic classification of objects in an array need not be not limited by the resources of working memory.

Belke et al. also tested for the effect of cognitive load on subjects' performance. Specifically, subjects were required to remember a series of digits during the presentation of the array (see Fig. 5). If subjects were using capacity-limited working memory mechanisms to rapidly identify the semantic content of each item in the array, one might expect their performance to be hindered by this task.

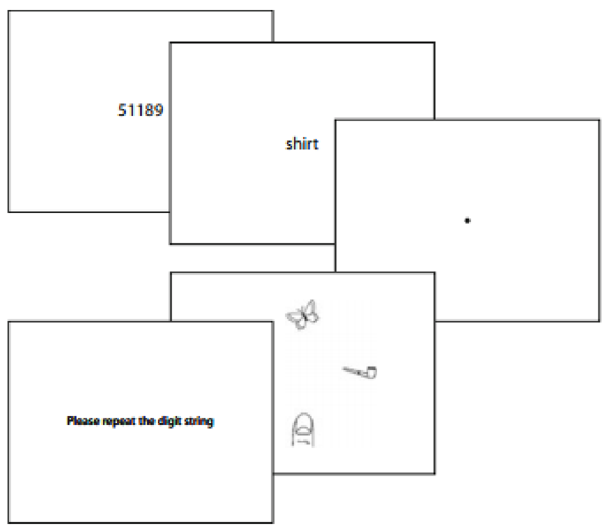

Figure 5 (Belke et al. 2008)

Subjects were briefly presented with five digits to hold in memory. They were then given a target description and told to search for an object matching that description in a presented array.

Instead, the results painted a more nuanced picture. Although cognitive load affected subjects' reaction times when *reporting* the presence or absence of a target, as well as broader features of their looking behavior (specifically causing subjects to linger longer on distractor items before moving on), it did not affect the initial looking behavior of subjects at all; subjects were just as likely to direct initial visual attention to target and distractor items in cognitive load cases as in all other variations of the experiment. This suggests that the mechanism by which the semantic identity of visual stimuli is retrieved (and which directs initial eye movements) does not rely on the same cognitive resources as working memory. As Belke et al. observe, their findings support the CSTM hypothesis and extend it, by "suggesting that there is parallel conceptual processing of visual stimuli prior to the first selection for attention" (Belke et al., 2008).[8]

As noted earlier, this review of the data is not intended to be exhaustive.[9] However, I hope I have given an illustrative sample of the varied paradigms that support the idea that there is short-term semantic processing of visually presented information outside of capacity-limited working memory, where this is not vulnerable to disruption by new visual input. The CSTM hypothesis explains this data by positing a high-capacity, fragile, short-term conceptual store. While many questions remain concerning CSTM – for example, whether the same mechanism is at work in different sensory modalities – I contend that we already have good reason for taking the CSTM hypothesis seriously.

3. CSTM in the philosophy of mind

If we take seriously the core claim of the CSTM hypothesis and suppose that there is indeed a stage of semantic memory intermediate between sensory processing and working memory, this by itself is of interest for theorists working on the mechanisms of human perception and cognition. However, I will now consider three specific philosophical debates in which CSTM might be of special relevance in empirically grounding existing accounts and opening up new theoretical spaces.

In applying CSTM to these debates, I will be working with the possibility that the contents of CSTM are sometimes conscious. It should be noted at the outset that it is very much

---

[8] The evidence for CSTM described above suggest a further puzzle for interpretation of earlier work on iconic memory: why, if there is rapid conceptual processing of items presented in an array, did subjects in Sperling's original paper fail to exhibit partial report superiority when cued to report on items in respect of their *semantic* identity (e.g., "letter" or "number")? One possibility is that subjects may have encoded the items in the array in respect of more specific semantic categories ("r" or "7") than those cued by Sperling ("letter" and "number"). Thus the number of characters that were reportable would be limited by subjects' ability to make rapid inferences about the determinable class the individual items belonged to. This would be a valuable area for further experimental work.

[9] Other important experiments explicitly engaging with the CSTM hypothesis include work on the attentional blink (Vogel, Luck, & Shapiro, 1998).

an open question whether this is the case. Potter herself is unsure on the point, and states that "the evidence… demonstrates that there is conceptual processing of material that is subsequently forgotten, [but] it does not tell us whether we were briefly conscious of that material, or whether the activation and selection occurred unconsciously." (Potter, 2012: 8). Indeed, as matters stand, it is likely to remain highly controversial whether any sort of short-term processing outside of working memory is conscious (Block, 2007): subjects' reports alone are unlikely to settle such matters, since what is ultimately reported by subjects will always be limited by what can be encoded in working memory. However, I think the hypothesis that CSTM is (at least sometimes) conscious is certainly worth exploring, and to the extent that it helps us to give a better account of perceptual experience we may have further reason to take it seriously. Note also that in what follows I will be providing only a brief survey of the relevant debates, and I will not make detailed proposals. Rather, my goal will be simply to outline how CSTM may allow for new perspectives and approaches in tackling these issues.

3.1 – CSTM and high-level phenomenology

One initial debate where CSTM could be of service concerns the claim that phenomenology non-derivatively represents high-level properties, also known as the Rich Content View (Bayne 2009, Siegel, 2010). This proposal holds that there are phenomenal properties associated not just with low-level features such as color and shape but a range of high-level features such as natural kinds, individuals, and causal relations. Thus it has been suggested that there may be distinctive kinds of phenomenal property associated with the perception of particular classes of objects: *looking like a bicycle*, for example, or *looking like a pine tree*.

Though the Rich Content View is far from uncontroversial (see, for example, Prinz, 2012a), many theorists are sympathetic to it, and it seems a potentially fruitful approach to understanding a range of phenomena, such as the apparent phenomenological shifts involved in expertise and perceptual learning (Siegel, 2010). However, one could agree with many of the examples of high-level phenomenology that defenders of the Rich Content View provide while denying that these cases are a matter of *perceptual* phenomenology. Consider, for example, the visual experience of someone who phenomenally represents a given tree as a Douglas Fir. One possibility is that this case, rather than being a perceptual effect, involves a subject making a rapid judgment with accompanying *cognitive* phenomenology.

While Siegel is open to the possibility of cognitive phenomenology, she rejects the idea that it can explain the full range of cases advanced as evidence of the Rich Content View. For example, in the following passage, she contrasts the (perhaps cognitive) phenomenology of carefully reading a sentence with merely glimpsing text in passing.

> Contrast this phenomenology [of reading] with that of being bombarded by pictures and captions on billboards along the highway. This seems a visual analog of the blare of a loud television, or a fellow passenger's inane cell-phone conversation: understanding the text on the billboard as you drive by isn't a deliberate affair; rather… it just happens… The advertisers would doubtless be happy if you lingered over every billboard's message, but no such event need occur I order for you to take in the semantic properties of the text as you whiz by. This suggests that the taking in can be merely sensory. (Siegel, 2010: 108)

Siegel is drawing a contrast here between our fleeting *perceptual* experience of semantic content that happens effortlessly and automatically and the more reflective *cognitive* awareness of such content that comes when we notice or pay close attention to things. Note that this latter presumably involves working memory, thus if her contrast is to be plausible we should aim to identify some further psychological mechanism besides working memory which might briefly encode incoming perceptual information in respect of high-level properties.

Here, then, is an initial place where CSTM might be put to good work. Recall that CSTM is fast, automatic, and fleeting, and seems to allow for the encoding of objects in respect of high level categories (Potter et al., 2014). Indeed, when Potter claims that CSTM constitutes the "basis for the unreflective understanding that is characteristic of everyday experience" (Potter, 2012), she may have in mind precisely the kind of case that Siegel is pointing to. Hence if the contents of CSTM were sometimes conscious, and contributed to the character of perceptual experience, we might find precisely the kind of high-level phenomenology that Siegel takes to be both ubiquitous and yet distinct from any kind of cognitive phenomenology associated with reflective and attentive thought.

Note also that purely sensory forms of memory are unlikely to provide a good explanation for the kind of content Siegel has in mind. As already noted, there is no evidence from Partial Report Paradigms that iconic memory is sensitive to high-level categories. There is also little neuroscientific support for the idea that high-level properties are engaged by early- or mid-level sensory areas of the brain. Thus Prinz (2012a), citing work by Vuielleumier et al. (2002), argues that the perceiving visually different objects as belonging to the same category involves the firing of neurons in frontal rather than sensory areas. More recent evidence from humans (Peelen & Caramazza, 2012) and primates (Perrodin et al., 2015) suggests that visual object recognition depends on activations very late in perceptual processing in the Anterior Temporal Cortex, an area associated with multi-modal or amodal semantic processing (Bonner & Price, 2013). As the authors of one study put it, “object representations in the anterior temporal lobes are abstracted away from perceptual properties, categorizing objects in semantically meaningful groups to support conceptual object knowledge” (Peelen & Caramazza, 2012).

Such data can be readily accommodated by a CSTM account of high level perceptual content, according to which high-level semantic content might arise very late in perceptual processing (perhaps combining information from multiple modalities). However, it is harder to reconcile with an account that claims high-level semantic processing occurs in iconic memory or fragile visual-short term memory, since these processes seem to be largely based in early visual areas (Sligte et al., 2011).

Note that nothing I have said counts against the more limited possibility (defended by Burge, 2010, and Carey, 2009) that sensory representations can represent non-basic properties like causal relations (though this is a highly contested point; see Shea, 2009). However, there is an good empirical case to made against the idea that early- or mid-level sensory processes could encode *all* of the very high level properties that Siegel takes to be present in perception, especially properties like “being a pine tree” or “being a CD player” that seem to rely on a wealth of fairly sophisticated knowledge of certain natural or artefactual kinds. Thus in light of the empirical resources currently available, CSTM or some similar intermediate mechanism seems to be the most plausible realization for the kind of high-level phenomenology that Siegel has in mind.

3.2 – CSTM and reportability

A second debate in which CSTM may be relevant concerns whether we sometimes consciously experience things while believing that we have not experienced them at all.[10] Several theories of conscious experience claim that whenever we consciously perceive something, we thereby *know* that we consciously perceived it, and can thus report it (and likewise, if we sincerely deny having seen something, then we did not consciously see it).

There are a variety of theoretical motivations for this view. Dehaene and Naccache, for example, offer a version of Global Workspace Theory that claims that consciousness consists in a form of cognitive access that automatically results in a state's becoming available for report (Dehaene & Naccache, 2001; Dehaene, 2014). Some higher-order thought theorists of consciousness have also defended a close relationship between consciousness and report (Rosenthal, 2005).

Whatever the merits of these theories, there is nonetheless some phenomenological motivation for resisting the conclusion that we must always be able to report the contents of our conscious experiences. For example, it is natural to think that we might see things without realizing we have done so. Imagine that I examine a multicolored object like a Buddhist mandala. Having scanned it, I am immediately asked whether I saw any red patches in the image, and I claim that I didn't. Upon repeating the task however, I realize that there were indeed a couple of red patches which I had looked at but not noticed. In cases like this, it seems at least an open possibility that we undergo brief conscious experiences that are not reflected in our reports.

There are a variety of ways in which one can respond to this sort of intuition (see Wu, 2014, for a review), but in practice, most participants in the debate about consciousness and reportability either adopt a 'sparse' theory of consciousness, challenging the intuition that there are ever conscious states that are not immediately available for report, or else claim that conscious experiences can be sustained by representations in sensory areas without access by any cognitive mechanisms (Block, 2007; Dretske, 1981; Lamme, 2010).

Both sorts of response have their downsides. On the one hand, sparse theories are arguably phenomenologically unconvincing, condensing all of perceptual experience into a fairly

[10] I have chosen to focus here on the debate about reportability rather than overflow (Block, 2007). The two debates are closely related, but some differences should be kept in mind. For example, one might allow that perceptual experience can be richer (or 'overflow') conscious thought while still being limited to states that subjects believe they experienced, even where they cannot provide details (that is, they must have "access to phenomenality"; see Balog, 2007).

narrow focus of attention.[11] For their part, many 'rich' views that allow for conscious experience without the active involvement of cognition may somewhat implausibly broaden the range of states that could potentially be conscious, leading to "panpsychic disaster" (Block, 2008). For example, if no cognitive processing whatsoever is required for conscious experience, it seems at possible that any sort of neural activity at all could be a potential realizer of conscious states, including activations in simple parts of the nervous system (such as the LGN or cerebellum) that are normally taken to be entirely subpersonal.

In this debate, the hypothesis that representations in CSTM may be briefly conscious could allow for a compromise of sorts between the two views. Recall that information in CSTM decays extremely rapidly, and subjects in Potters' experiments were typically unable to recall presented pictures or words if cued more than a second or two after presentation. Note also that *noticing* an object for the preparing report or deliberate action (as opposed to merely encoding semantic information about it) requires the involvement of working memory. Hence if CSTM is sometimes conscious, then there may be conscious conceptual representations in CSTM that subjects fail to notice and therefore might deny having seen.[12]

This position would allow us to hold that conscious experience really is sometimes richer than suggested by subjects' reports alone, and can thus more easily accommodate the phenomenological data without recourse to debunking accounts. It does this, however, without committing to the more extreme view (offered by Block and Lamme) that consciousness can be an entirely non-cognitive matter. One might hold, for example, that in order to become conscious, information must be at least be subject to some kind of initial semantic processing of the kind found in CSTM, even if it is only fleetingly available to working memory. This would entail that any conscious state must be integrated at least briefly into a subject's cognitive economy, thus ruling out many of the more extreme examples raised against Block and Lamme, such as wholly inaccessible conscious states in early vision or other presumably subpersonal

[11] There are, of course, numerous ways in which a sparse theorist can attempt to square their account with the phenomenological data. Many of the proposals have attempted to explain away the seemingly large capacity of phenomenal consciousness in terms of 'generic' or 'fragmentary' phenomenology, arguing that our seemingly rich visual world is compatible with conscious experience being based in strictly capacity-limited mechanisms like working memory (Grush, 2007; Kouider, Gardelle, Sackur, & Dupoux, 2010; Stazicker, 2011). A quite different sort of reply is given by Phillips (2011) who suggests that we can understand partial report paradigms without overflow by rejecting the picture, typically assumed in the debate, of how experiences unfold over time.

[12] This may also be one way to account for subjects' experience in attentional blink paradigms. See Vogel, Luck, & Shapiro (1998) for such an account explicitly couched in terms of CSTM.

systems.[13] This is a tentative proposal, of course, but points to another debate where CSTM opens up new theoretical possibilities.

3.3 – CSTM as a bridge between perception and cognition

Thus far, I have referred to CSTM somewhat loosely as a form of perceptual memory, but it exhibits a strange mix of the usual features that distinguish perception from cognition. For example, while it operates extremely rapidly and seemingly involuntarily, and is not automatically available for report, it nonetheless encodes information in a conceptual rather than purely sensory format. While there is no universal consensus on how to distinguish perception from cognition, one way of carving up the two (Dretske, 1981) claims that perception is distinguished by its nonconceptual analogue content, while cognition involves propositional attitudes such as believing, judging, and thinking. But while the conceptual format of CSTM might seem to place it on the cognitive side of this divide, it does not seem to support any of our ordinary propositional attitude ascriptions: it would seem strange to say that a subject who fleetingly conceptually classifies an image in CSTM as being a picture of a wedding thereby *thinks* or *believes* that the picture is of a wedding, for example.

Potter herself has claimed that CSTM shows that "perception is continuous with cognition" (Potter, 2012:9). However, another option would be to claim that CSTM should be considered a distinct process in its own right – a third psychological kind, perhaps. Coming after perception, but prior to highly capacity-limited cognitive processes, CSTM might constitute a kind of buffer in which sensory inputs first get conceptualized in respect of their semantic category. These conceptualized contents could then be made available to working memory systems and accessed selectively depending on a subject's goals and interests at a given moment.[14]

This is just one speculative suggestion, but simply by raising the possibility of a stage of

[13] Another way to rule out such states is to insist that only sensory states that are *accessible* to a subject's cognitive systems can be conscious. This sort of view is held by Prinz (2012b) and Carruthers (2011). However, these theories face the challenge of explaining how a state's merely being *disposed* to be accessed could be responsible for the occurrent phenomenon of consciousness. A CSTM view, by contrast, can appeal to an occurrent process, namely active encoding in CSTM, to explain why these states are consciously experienced.

[14] If CSTM is indeed distinct from both perception and cognition, then the history of philosophy and psychology might furnish us with an apt name for it in the notion of *apperception*. As James (1900) puts it, "we never get an experience that remains for us completely nondescript: it always reminds of something similar in quality, or of some context that might have surrounded it before, and which it now in some way suggests… We conceive the impression in some definite way. … This way of taking in the object is the process of apperception."

processing intermediate between perception and cognition, CSTM may be of relevance to debates that trade on the nature of the boundary between the two. Of particular note in this regard are the various debates concerning the extent to which cognitive processes can affect (or penetrate) perceptual processing. A copious amount of experimental data has been produced in support of the idea that there are 'top-down' effects on perception, which in turn has been taken to suggest that one's thoughts, beliefs, and desires can significantly alter the way the world appears. For example, one result showed that subjects instructed to throw a *heavy* ball at a target judge the distance to the target to be greater than subjects throwing a *lighter* ball, suggesting that their beliefs about the ball may have influenced the apparent distance to the target (Witt et al., 2004).

However, such conclusions are increasingly being contested on methodological and theoretical grounds. Thus Firestone and Scholl (2015) claim that many such results fail to properly distinguish *perception* from *judgment*, claiming that, in many cases, experimentalists' results can be interpreted purely in terms of post-perceptual *cognitive* effects rather than as involving effects on perception. In this spirit, Firestone and Scholl claim that it is important to distinguish perception from judgment carefully when assessing whether top-down effects are present, and to avoid blurring the lines between the two. As they warn us, "many papers in this literature advert to effects on "perceptual judgment"… which can only invite confusion about this foundational distinction."

However, if CSTM is a distinct psychological process operative between perception and judgment, then new theoretical possibilities present themselves in the cognitive penetration debate. For example, it might be that the operations of CSTM can themselves be affected by an individual's beliefs and desires, even if there is no direct cognitive influence on low- and mid-level sensory processing (thus also respecting the claim that early vision is encapsulated; see Pylyshyn, 1999). If this were the case, then it would mean that while the strictly sensory elements of perception (such as basic shades and contours) were not directly penetrated by cognition, the conceptual contents of perception (such as apparent color category and shape category) might indeed by thus penetrated.

What this possibility would mean for the broader cognitive penetration debate would depend on broader theoretical commitments. A theorist sympathetic to the Rich Content View discussed earlier might claim that the way that sensory content is conceptualized might indeed affect the overall phenomenal character of perceptual experience, albeit in respect of respect of higher level conceptual properties rather than low-level sensory ones. But even for theorists who deny that high-level contents directly affect the character of perceptual experience in this way, the above proposal still offers a new theoretical route to understanding apparent cases of top-down effects on perception.

Of course, the account suggested is again just an outline: further empirical and theoretical work would be helpful to explore, for example, whether the kind of rapid and automatic processes of perceptual categorization seen in Potter's work can be influenced by subjects' broader beliefs and desires. However, I take it that this again constitutes a debate in which CSTM may serve to create new theoretical possibilities and suggest further research.

4. **Conclusion**

My central goal in this paper has been to examine evidence for a proposed distinctive form of memory namely CSTM, and to explore some ways in which it may illuminate some important questions in philosophy and cognitive science and open up new theoretical approaches. Whether or not these options turn out to be plausible will require further empirical work, but it is possible that in some of these debates, CSTM may have been a piece of the mind that has been missing from the theoretical puzzle.

REFERENCES

Atkinson, R., & Shiffrin, R., 1968: Human Memory: A Proposed System and its Control Processes. *Psychology of Learning and Motivation*, 89-195.

Averbach, E., & Coriell, A. S., 1961: Short-Term Memory in Vision. *Bell System Technical Journal*, 401, 309-328.

Baddeley, A.D. (2000). The episodic buffer: A new component of working memory? *Trends in Cognitive Sciences*, 4(11), 417-423.

Baddeley, A. D., Lewis, V., & Vallar, G., 1984. Exploring the articulatory loop. *The Quarterly Journal of Experimental Psychology Section A, 36*2, 233-252.

Baddeley, A. D., Thomson, N., & Buchanan, M. 1975. Word length and the structure of short-term memory. *Journal of Verbal Learning and Verbal Behavior, 14*6, 575-589. doi:10.1016/s0022-53717580045-4

Balog, K. 2007: Access to phenomenality: A necessary condition of phenomenality? *Behavioral and Brain Sciences*, 305-6.

Bayne, T., 2009: Perception And The Reach Of Phenomenal Content. *The Philosophical Quarterly*, 59236, 385-404.

Belke, E., Humphreys, G., Watson, D., Meyer, A. and Telling, A., 2008: Top-down effects ofsemantic knowledge in visual search are modulated by cognitive but not perceptual load. *Perception & Psychophysics*, 70 8, 1444 – 1458.

Block, N., 2007: Consciousness, accessibility, and the mesh between psychology and neuroscience. *Behavioral and Brain Sciences*, 305-6.

Block, N., 2008. Consciousness and Cognitive Access. *Proceedings of the Aristotelian Society Hardback, 108*, 289-317.

Bonner, M. F., & Price, A. R., 2013. Where Is the Anterior Temporal Lobe and What Does It Do? *Journal of Neuroscience*, *33*10, 4213-4215.

Brady, T. F., Störmer, V. S., & Alvarez, G. A. (2016). Working memory is not fixed-capacity: More active storage capacity for real-world objects than for simple stimuli. *Proceedings of the National Academy of Sciences* USA, 113(27), 7459-7464.

Burge, T., 2010. *Origins of objectivity*. Oxford: Oxford University Press.

Carey, S., 2009. *The origin of concepts*. Oxford: Oxford University Press.

Carruthers, P., 2011: *The opacity of mind: An integrative theory of self-knowledge*. Oxford: Oxford University Press.

Cowan, N. (1988). Evolving conceptions of memory storage, selective attention, and their mutual constraints within the human information processing system. *Psychological Bulletin*, 104, 163–191.

Cowan, N., 2001. The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences 24*1, 87-114.

Darwin, C. J., Turvey, M. T., & Crowder, R. G. 1972: An auditory analogue of the Sperling partial report procedure: Evidence for brief auditory storage. *Cognitive Psychology*, 32, 255-267.

Dehaene, S., 2014: *Consciousness and the Brain: Deciphering How the Brain Codes Our Thoughts*. Viking Press, 2014.

Dehaene, S., & Naccache, L., 2001: Towards a cognitive neuroscience of consciousness: Basic evidence and a workspace framework. *Cognition*, 79:1-37.

Dretske, F., 1981: *Knowledge and the Flow of Information*. MIT Press.

Firestone, C., & Scholl, B. J., 2015. Cognition does not affect perception: Evaluating the evidence for 'top-down' effects. *Behavioral and Brain Sciences* 1-77.

Fougnie, D. 2008: The Relationship between Attention and Working Memory. In Johansen, N. B. (Ed.). *New research on short-term memory*. New York: Nova Biomedical Books.

Grush, R., 2007: A plug for generic phenomenology. *Behavioral and Brain Sciences*, 305-6.

Harris, J., Miniussi, C., Harris, I., & Diamond, M., 2002. Transient storage of a tactile memory trace in primary somatosensory cortex. Journal of Neuroscience, 22, 8720–8725

James, W., 1900. *Talks to teachers and students*. New York: Henry Holt.

Kouider, S., Gardelle, V. D., Sackur, J., & Dupoux, E., 2010: How rich is consciousness? The partial awareness hypothesis. *Trends in Cognitive Sciences*, 147, 301-307.

Lamme, V. A., 2010. How neuroscience will change our view on consciousness. *Cognitive Neuroscience, 1*3, 204-220.

Landman, R., Spekreijse, H., & Lamme, V. A. F., 2003: Large capacity storage of integrated objects before change blindness. *Vision Research*, 43, 149–164.

Luck, S. J., & Vogel, E. K., 1997: The capacity of visual working memory for features and conjunctions. *Nature*, 390, 279-281.

Ma, W. J., Husain, M., & Bays, P. M., 2014. Changing concepts of working memory. *Nature Neuroscience*, *17*3, 347-356.

Maguire, J. F., & Howe, P. D. (2016). Failure to detect meaning in RSVP at 27 ms per picture. *Attention, Perception, & Psychophysics* 78(5), 1405-1413.

Miller, G. A., 1956: The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review*, 632, 81-97.

Moores E., Laiti L., Chelazzi L., 2003 : Associative knowledge controls deployment of visual selective attention. *Nature Neuroscience* 6, 182–18.

Peelen, M. V., & Caramazza, A., 2012. Conceptual Object Representations in Human Anterior Temporal Cortex. *Journal of Neuroscience, 32*45, 15728-15736.

Perrodin, C., Kayser, C., Abel, T. J., Logothetis, N. K., & Petkov, C. I., 2015. Who is That? Brain Networks and Mechanisms for Identifying Individuals. *Trends in Cognitive Sciences*, *19*12, 783-796.

Phillips, I. B., 2011: Perception and Iconic Memory: What Sperling Doesn't Show. *Mind & Language*, 264, 381-411.

Potter, M., 2009: Conceptual short term memory. *Scholarpedia*, 52, 3334.

Potter, M. C., 2012: Conceptual Short Term Memory in Perception and Thought. *Frontiers in Psychology,* 3:113.

Potter, M.C., Kroll, J.F., & Harris, C. 1980: Comprehension and memory in rapid sequential reading. In R. Nickerson Ed., *Attention and Performance VIII* pp. 395-418. Hillsdale, NJ: Erlbaum.

Potter, M.C., Stiefbold, D., & Moryadas, A. 1998: *Word selection in reading sentences: Preceding versus following contexts. Journal of Experimental Psychology: Learning, Memory, and Cognition*, 24, 68-100.

Potter, M. C., Staub, A., & O'Connor, D. H., 2004: Pictorial and conceptual representation of glimpsed pictures. *Journal of Experimental Psychology: Human Perception and Performance*, 30, 478-489.

Potter, M.C., Wyble, B., Hagmann, C.E., & McCourt, E.S., 2014: Detecting meaning in RSVP at 13 ms per picture. *Attention, Perception, & Psychophysics*, 762, 270-279.

Prinz, J., 2012a. Siegel’s get rich quick scheme. *Stud Philosophical Studies,163*3, 827-835.

Prinz, J., 2012b: *The conscious brain: How attention engenders experience*. Oxford: Oxford University Press.

Pylyshyn, Z., 1999. Is vision continuous with cognition?: The case for cognitive impenetrability of visual perception. *Behavioral and Brain Sciences, 22*(03). doi:10.1017/s0140525x99002022

Rosenthal, D., 2005: *Consciousness and Mind*, Oxford: Clarendon Press.

Salmela, V. R., Moisala, M., & Alho, K., 2014: Working memory resources are shared across sensory modalities. *Attention, Perception, & Psychophysics*, 767, 1962-1974.

Saults, J.S., & Cowan, N.,, 2007: A central capacity limit to the simultaneous storage of visual and auditory arrays in working memory. *Journal of Experimental Psychology: General*, 1364, 663-684.

Shea, N. (2009). New concepts can be learned. *Biology & Philosophy 26*(1), 129-139.

Siegel, S., 2010: *The contents of visual experience.* New York: Oxford University Press.

Sligte, I. G., Scholte, H. S., & Lamme, V. A., 2008 : Are There Multiple Visual Short-Term Memory Stores? *PLoS ONE*, 32.

Sligte, I. G., Wokke, M. E., Tesselaar, J. P., Scholte, H. S., & Lamme, V. A., 2011: Magnetic stimulation of the dorsolateral prefrontal cortex dissociates fragile visual short-term memory from visual working memory. *Neuropsychologia*, 496, 1578-1588.

Sperling, G., 1960: The information available in brief visual presentations. *Psychological Monographs: General and Applied*, 7411, 1-29.

Stazicker, J., 2011: Attention, Visual Consciousness and Indeterminacy. *Mind & Language*, 262, 156-184.

Stein, T., Kaiser, D., Hesselmann, G. 2016: Can working memory be non-conscious? *Neuroscience of Consciousness*, 2016 1, 1-3.

Vogel, E. K., Luck, S. J. & Shapiro, K. L. (1998) Electrophysiological evidence for a postperceptual locus of suppression during the attentional blink. *Journal of Experimental Psychology: Human Perception and Performance* 24:1656–74.

Von Wright, J., 1970: On selection in visual immediate memory. *Acta Psychologica*, 33, 280-292.

Vuilleumier, P., Henson, R. N., Driver, J., & Dolan, R. J., 2002. Multiple levels of visual object constancy revealed by event-related fMRI of repetition priming. *Nature Neuroscience*, 5, 491–499.

Witt, J. K., Proffitt, D. R., & Epstein, W., 2004. Perceiving distance: A role of effort and intent. Perception, 33, 577–590.

Wu, W., 2014: *Attention*, Routledge.